

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model
Run on: May 25, 2005, 17:08:14 ; Search time 499 Seconds
(without alignments)
912.871 Million cell updates/sec

Title: US-10-032-585-7068
Perfect score: 1997
Sequence: 1 MSQINPLGSDIAKLIP.....QGLVDANGNPIPGAPGAL 390

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 6959266 seqs, 1168006243 residues

Total number of hits satisfying chosen parameters: 6959266

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: /cgn2_6/ptodata/1/paa/PCTUS COMB.pcp.*
 - 2: /cgn2_6/ptodata/1/paa/US06 COMB.pcp.*
 - 3: /cgn2_6/ptodata/1/paa/US07 COMB.pcp.*
 - 4: /cgn2_6/ptodata/1/paa/US080 COMB.pcp.*
 - 5: /cgn2_6/ptodata/1/paa/US081 COMB.pcp.*
 - 6: /cgn2_6/ptodata/1/paa/US082 COMB.pcp.*
 - 7: /cgn2_6/ptodata/1/paa/US083 COMB.pcp.*
 - 8: /cgn2_6/ptodata/1/paa/US084 COMB.pcp.*
 - 9: /cgn2_6/ptodata/1/paa/US085 COMB.pcp.*
 - 10: /cgn2_6/ptodata/1/paa/US086 COMB.pcp.*
 - 11: /cgn2_6/ptodata/1/paa/US087 COMB.pcp.*
 - 12: /cgn2_6/ptodata/1/paa/US088 COMB.pcp.*
 - 13: /cgn2_6/ptodata/1/paa/US089 COMB.pcp.*
 - 14: /cgn2_6/ptodata/1/paa/US090 COMB.pcp.*
 - 15: /cgn2_6/ptodata/1/paa/US091 COMB.pcp.*
 - 16: /cgn2_6/ptodata/1/paa/US092 COMB.pcp.*
 - 17: /cgn2_6/ptodata/1/paa/US093 COMB.pcp.*
 - 18: /cgn2_6/ptodata/1/paa/US094 COMB.pcp.*
 - 19: /cgn2_6/ptodata/1/paa/US095 COMB.pcp.*
 - 20: /cgn2_6/ptodata/1/paa/US096 COMB.pcp.*
 - 21: /cgn2_6/ptodata/1/paa/US097A COMB.pcp.*
 - 22: /cgn2_6/ptodata/1/paa/US097B COMB.pcp.*
 - 23: /cgn2_6/ptodata/1/paa/US098 COMB.pcp.*
 - 24: /cgn2_6/ptodata/1/paa/US099A COMB.pcp.*
 - 25: /cgn2_6/ptodata/1/paa/US099B COMB.pcp.*
 - 26: /cgn2_6/ptodata/1/paa/US100 COMB.pcp.*
 - 27: /cgn2_6/ptodata/1/paa/US101 COMB.pcp.*
 - 28: /cgn2_6/ptodata/1/paa/US102 COMB.pcp.*
 - 29: /cgn2_6/ptodata/1/paa/US103 COMB.pcp.*
 - 30: /cgn2_6/ptodata/1/paa/US104 COMB.pcp.*
 - 31: /cgn2_6/ptodata/1/paa/US105 COMB.pcp.*
 - 32: /cgn2_6/ptodata/1/paa/US106 COMB.pcp.*
 - 33: /cgn2_6/ptodata/1/paa/US107 COMB.pcp.*
 - 34: /cgn2_6/ptodata/1/paa/US108 COMB.pcp.*
 - 35: /cgn2_6/ptodata/1/paa/US109 COMB.pcp.*
 - 36: /cgn2_6/ptodata/1/paa/US110 COMB.pcp.*
 - 37: /cgn2_6/ptodata/1/paa/US60 COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1997	100.0	390	1	PCT-US02-03987-15012	Sequence 15012, A
2	1997	100.0	390	26	US-10-032-585-7068	Sequence 7068, A
3	1997	100.0	390	26	US-10-072-851-15012A	Sequence 15012, A
4	1997	100.0	390	37	US-60-259-128-4972	Sequence 4972, A
5	1997	100.0	390	37	US-60-314-050-7068	Sequence 7068, A
6	1997	100.0	414	27	US-10-179-131-7485	Sequence 7485, A
7	1976	98.9	419	32	US-60-603-113-17623	Sequence 17623, A
8	1976	98.9	419	37	US-60-096-409-17623	Sequence 17623, A
9	1214.5	60.8	387	29	US-10-326-956-3117	Sequence 3117, A
10	745	37.3	355	1	PCT-US02-13142-3293	Sequence 3293, A
11	745	37.3	355	1	PCT-US04-01099-3380	Sequence 3380, A
12	745	37.3	355	27	US-10-128-714-3293	Sequence 3293, A
13	745	37.3	355	33	US-10-760-089-3380	Sequence 3380, A
14	745	37.3	355	37	US-60-316-362-3293	Sequence 3293, A
15	745	37.3	413	20	US-09-675-7848-9454	Sequence 9454, A
16	695	34.8	336	1	PCT-US02-13142-8293	Sequence 8293, A
17	695	34.8	336	27	US-10-128-714-8293	Sequence 8293, A
18	670	33.6	338	37	US-60-285-697-2	Sequence 2, Appli
19	670	33.6	338	37	US-60-295-890-2	Sequence 2, Appli
20	542	27.1	375	1	PCT-US02-40225-3129	Sequence 3129, A
21	542	27.1	375	29	US-10-320-797-3129	Sequence 3129, A
22	542	27.1	375	37	US-60-341-261-3129	Sequence 3129, A
23	320.5	16.0	344	20	US-09-614-150-4386	Sequence 4386, A
24	320.5	16.0	344	20	US-09-614-150A-4386	Sequence 4386, A
25	320.5	16.0	344	37	US-60-167-217-4469	Sequence 4469, A
26	320.5	16.0	344	37	US-60-173-464-3662	Sequence 3662, A
27	320.5	16.0	344	37	US-60-191-637-4401	Sequence 4401, A
28	320.5	16.0	344	37	US-60-191-681-3504	Sequence 3504, A
29	290	14.5	340	20	US-09-614-150A-2103	Sequence 2103, A
30	290	14.5	340	20	US-09-614-150A-2103	Sequence 2103, A
31	290	14.5	340	24	US-09-949-029-50	Sequence 50, Appl
32	290	14.5	340	37	US-60-191-637-2109	Sequence 2109, A
33	290	14.5	340	37	US-60-191-681-1686	Sequence 1686, A
34	263	13.2	361	1	PCT-US04-24424-1055	Sequence 1055, A
35	263	13.2	361	27	US-10-170-205E-21742	Sequence 21742, A
36	263	13.2	361	30	US-10-408-765-1604	Sequence 1604, A
37	263	13.2	361	30	US-10-408-765A-1604	Sequence 1604, A
38	263	13.2	361	37	US-60-389-987-1604	Sequence 1604, A
39	263	13.2	361	37	US-60-412-418-1604	Sequence 1604, A
40	263	13.2	361	37	US-60-452-680-19817	Sequence 19817, A
41	263	13.2	361	37	US-60-453-050-11927	Sequence 11927, A
42	263	13.2	361	37	US-60-453-135-11927	Sequence 11927, A
43	263	13.2	361	37	US-60-466-412-11927	Sequence 11927, A
44	263	13.2	361	37	US-60-490-890-1055	Sequence 1055, A
45	263	13.2	389	1	PCT-US01-08631-52270	Sequence 52270, A

ALIGNMENTS

RESULT 1

PCT-US02-03987-15012
; Sequence 15012, Application PC/TUS0203987
; GENERAL INFORMATION:
; APPLICANT: Elittra Pharmaceuticals, Inc.
; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits
; TITLE OF INVENTION: Proliferation
; FILE REFERENCE: ELITRA-028VPC
; CURRENT APPLICATION NUMBER: PCT/US02/03987
; PRIOR FILING DATE: 2002-02-02
; CURRENT FILING DATE: 2002-02-02
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15012
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Candida albicans
PCT-US02-03987-15012

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2005, 17:05:48 ; Search time 43 Seconds
(without alignments)

677.050 Million cell updates/sec

Title: US-10-032-585-7068

Perfect score: 1997

Sequence: 1 MSQIINPLGSTDIAKLSIP.....QGLVDANGNPIGAPAPGAL 390

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5A COMB.pap.*

2: /cgn2_6/ptodata/1/iaa/5B COMB.pap.*

3: /cgn2_6/ptodata/1/iaa/6A COMB.pap.*

4: /cgn2_6/ptodata/1/iaa/6B COMB.pap.*

5: /cgn2_6/ptodata/1/iaa/6B COMB.pap.*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1976	98.9	419	4	US-09-248-796A-17623
2	329.5	16.5	453	4	US-09-270-767-43715
3	125.5	6.3	204	4	US-09-270-767-59105
4	108	5.4	605	4	US-09-489-039A-13002
5	102	5.1	1037	4	US-09-543-681A-7677
6	101	5.1	737	4	US-09-071-035-460
7	101	5.1	2032	4	US-09-071-035-458
8	101	5.1	2032	4	US-09-071-035-462
9	101	5.1	2032	4	US-09-071-035-466
10	101	5.1	2054	4	US-09-134-000C-6612
11	100.5	5.0	609	4	US-09-252-991A-28399
12	99	5.0	509	4	US-09-583-110-5241
13	99	5.0	512	4	US-09-107-433-5122
14	98.5	4.9	1935	4	US-09-538-092-916
15	98.5	4.9	1944	4	US-09-949-016-10929
16	98	4.9	671	4	US-09-198-452A-468
17	98	4.9	947	4	US-09-438-185A-447
18	97.5	4.9	4563	4	US-09-108-006C-1
19	95.5	4.8	840	4	US-09-079-030-214
20	95.5	4.8	863	2	US-08-666-271-2
21	95.5	4.8	4536	4	US-09-180-422B-27
22	95.5	4.8	4536	4	US-09-079-030-1
23	95.5	4.8	4563	4	US-09-538-092-842
24	95	4.8	477	4	US-09-328-352-7765
25	95	4.8	653	4	US-09-438-185A-730
26	95	4.8	870	2	US-08-266-311-1
27	95	4.8	870	2	US-08-467-527A-1

28	95	4.8	870	3	US-08-467-528-1	Sequence 1, Appli
29	95	4.8	1646	4	US-09-583-110-4667	Sequence 4667, Ap
30	95	4.8	1653	4	US-09-107-433-4155	Sequence 4155, Ap
31	94.5	4.7	491	3	US-08-942-012B-27	Sequence 27, Appli
32	94	4.7	722	1	US-08-445-050-7	Sequence 7, Appli
33	94	4.7	722	2	US-08-204-691-7	Sequence 7, Appli
34	93.5	4.7	397	4	US-09-270-767-45798	Sequence 45798, A
35	93.5	4.7	1454	4	US-09-328-352-5793	Sequence 5793, Ap
36	93	4.7	1095	4	US-08-851-567B-34	Sequence 34, Appli
37	92.5	4.6	283	1	US-08-726-227-1	Sequence 1, Appli
38	92.5	4.6	283	3	US-08-826-246-6	Sequence 6, Appli
39	92.5	4.6	283	3	US-08-944-495-6	Sequence 6, Appli
40	92.5	4.6	283	3	US-09-126-640-9	Sequence 9, Appli
41	92.5	4.6	283	3	US-08-925-588-6	Sequence 6, Appli
42	92.5	4.6	283	3	US-09-288-292A-9	Sequence 9, Appli
43	92.5	4.6	283	4	US-09-372-044-6	Sequence 6, Appli
44	92.5	4.6	283	4	US-08-825-486-6	Sequence 6, Appli
45	92.5	4.6	283	4	US-08-826-248-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1

US-09-248-796A-17623

; Sequence 17623, Application US/09248796A

; Patent No. 6747137

; GENERAL INFORMATION:

; APPLICANT: Keith Weinstein et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA

; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.132

; CURRENT APPLICATION NUMBER: US/09/248,796A

; CURRENT FILING DATE: 1999-02-12

; PRIOR APPLICATION NUMBER: US 60/074,725

; PRIOR FILING DATE: 1998-02-13

; PRIOR APPLICATION NUMBER: US 60/096,409

; PRIOR FILING DATE: 1998-08-13

; NUMBER OF SEQ ID NOS: 28208

; SEQ ID NO 17623

; LENGTH: 419

; TYPE: PRT

; ORGANISM: Candida albicans

; US-09-248-796A-17623

Query Match 98.9%; Score 1976; DB 4; Length 419;

Best Local Similarity 99.5%; Pred. No. 9.8e-197;

Matches 387; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	1	MSQIINPLGSTDIAKLSIPTLPQVTTEPKQGLWSSNPVFSYINDVYITINEHRKSLG	60
Db	25	MSQIINPLGSTDIAKLSIPTLPQVTTEPKQGLWSSNPVFSYINDVYITINEHRKSLG	84
Qy	61	LTPNGTIENLKEVARDVFLGQYFTGLRADLNKAFSMMPAFOTSHTLSTGNSVLPAYAF	120
Db	85	LTPNGTIENLKEVARDVFLGQYFTGLRADLNKAFSMMPAFOTSHTLSTGNSVLPAYAF	144
Qy	121	SALYATDDYFLOGNINDLSFSGRINYGWDSKNSKVTLOLAHQPSMIQLEQDYQANDC	180
Db	145	SALYATDDYFLOGNINDLSFSGRINYGWDSKNSKVTLOLAHQPSMIQLEQDYQANDC	204
Qy	181	SINVKTNPNFLSGNFSGVVGSILQSLSKLAVGLTMYSKOPLAPPTAVSVARYN	240
Db	205	SINVKTNPNFLSGNFSGVVGSILQSLSKLAVGLTMYSKOPLAPPTAVSVARYN	264
Qy	241	AGNWIASAQLOAQGALIASFWRKVTDKVBAAGLETQVAATMKQVADPLMGVGFPEVIEQQT	300
Db	265	AGNWIASAQLOAQGALIASFWRKVTDKVBAAGLETQVAATMKQVADPLMGVGFPEVIEQQT	324
Qy	301	TIGAKYRTAVFRGQLDSKIGKISAFLEKRIPTVSIILFSGEIDQFKNTSLGLGLOPEA	360
Db	325	TIGAKYRTAVFRGQLDSKIGKISAFLEKRIPTVSIILFSGEIDQFKNTSLGLGLOPEA	384

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2005, 16:56:27 ; Search time 163 Seconds
(without alignments)
925.378 Million cell updates/sec

Title: US-10-032-585-7068
Perfect score: 1997
Sequence: 1 MSQINPLGSTDIAKLSP.....QGLVDANGNPIPGAPGAL 390

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04.*
1: Geneseq_1980s.*
2: Geneseq_1990s.*
3: Geneseq_2000s.*
4: Geneseq_2001s.*
5: Geneseq_2002s.*
6: Geneseq_2003as.*
7: Geneseq_2003bs.*
8: Geneseq_2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1997	100.0	390	5	ABP73231 Candida a
2	1214.5	60.8	387	6	ABR53859 Protein s
3	1214.5	60.8	387	7	ADK64914 Disease t
4	745	37.3	355	6	ABJ25635 Aspergill
5	745	37.3	355	8	ADR86330 Aspergill
6	695	34.8	336	6	ABJ26235 Aspergill
7	542	27.1	375	7	ADB70085 C. neofor
8	320.5	16.0	344	4	ABBS3198
9	290	14.5	340	4	ABBS58437 Drosophil
10	290	14.5	340	7	ADJ37933 D melanog
11	263	13.2	361	7	ADJ69798 Human hea
12	263	13.2	361	8	ABM82134 Tumour-as
13	263	13.2	361	8	ADP56258 Human PRO
14	263	13.2	361	8	ADP25360 PRO poly
15	263	13.2	389	4	ABG21911 Novel hum
16	258	12.9	314	7	ADJ70870 Human hea
17	251.5	12.6	349	6	ABR44273 Novel hum
18	242.5	12.1	309	3	ABG1737 Arabidops
19	241.5	12.1	309	3	ABG20918 Arabidops
20	241.5	12.1	332	3	ABG20917 Arabidops
21	239	12.0	308	4	ABG93239 Human pro
22	239	12.0	308	6	ABG99924 Human nov
23	239	12.0	318	7	ADI21663 Novel hum
24	234.5	11.7	267	3	ABG1738 Arabidops
25	233.5	11.7	254	2	AAW82285 Natural k

26	233.5	11.7	267	3	ABG20919 Arabidops
27	217.5	10.9	324	3	ABW58832 Breast an
28	203	10.2	243	2	AAW82283 Tumour su
29	199.5	10.0	238	3	ABG11739 Arabidops
30	188.5	9.4	112	4	ABM83464 Human imm
31	175.5	8.8	257	5	ABG99082 Apolipop
32	171.5	8.6	276	4	ABG21910 Novel hum
33	160.5	8.0	190	4	ABG95169 Human pro
34	133.5	6.2	1333	6	ABU27622 Protein e
35	116.5	5.8	320	4	ABW11307 Drosophil
36	113.5	5.7	1122	6	ABU14859 Protein e
37	111	5.6	927	6	ABG36312 Moraxella
38	110.5	5.5	321	4	ABG19571 Novel hum
39	108	5.4	605	7	ABO66485 Klebsiell
40	107.5	5.4	276	3	ABG05860 Arabidops
41	107	5.4	927	4	ABG62110 M. bovis
42	106.5	5.3	276	3	ABG53428 Arabidops
43	106.5	5.3	276	5	ABG92164 Herbicida
44	106	5.3	275	7	ABM74317 DNA clone
45	105.5	5.3	1935	7	ADD45207 Rat Prote

ALIGNMENTS

RESULT 1
ABP73231
ID ABP73231 standard; protein; 390 AA.
XX AC ABP73231;
XX AC ABP73231;
DT 30-JAN-2003 (first entry)
XX
DE Candida albicans essential protein SEQ ID NO 7068.
XX
KW Fungus; yeast; tetracyclin; promoter; GRACE strain; biosynthesis;
KW signal transduction; DNA replication; cell division; growth;
KW proliferation; Candida albicans; fungicide; antifungal.
XX
OS Candida albicans.
XX
BN W0200253728-A2.
XX
PD 11-JUL-2002.
XX
PF 26-DEC-2001; 2001WO-US049486.
XX
PR 29-DEC-2000; 2000US-0259128P.
PR 20-FEB-2001; 2001US-00792024.
PR 22-AUG-2001; 2001US-0314050P.
(ELIT-) ELITRA PHARM INC.
PI Roemer T, Jiang B, Boone C, Bussey H, Ohlsen KL;
XX WPI; 2002-566694/60.
DR N-PSDB; ABZ31781.
XX
PT Constructing strains for identifying gene products as effective targets
PT for therapeutic intervention, by inactivating in the strain one allele of
PT a gene and placing other allele of the gene under conditional expression.
XX
PS Claim 44; SEQ ID NO 7068; 167pp + Sequence Listing; English.
XX
CC The invention relates to constructing (M1) a strain of diploid fungal
CC cells in which both alleles of a gene are modified, comprising modifying
CC one allele by insertion or replacement by a cassette having an
CC expressible selectable marker and modifying other allele by
CC recombination, of a promoter replacement fragment with a heterologous
CC promoter, so that expression of the second allele is regulated by the
CC promoter. (M1) is useful for constructing a strain of diploid fungal
CC cells in which both alleles of a gene are modified. The diploid fungal
CC cells having both alleles modified are useful for identifying a gene that